MINAS-a database of Metal Ions in Nucleic AcidS

Schnabl, J ; Suter, P ; Sigel, Roland K O

Abstract: Correctly folded into the respective native 3D structure, RNA and DNA are responsible for uncountable key functions in any viable organism. In order to exert their function, metal ion cofactors are closely involved in folding, structure formation and, e.g. in ribozymes, also the catalytic mechanism. The database MINAS, Metal Ions in Nucleic AcidS (http://www.minas.uzh.ch), compiles the detailed information on innersphere, outersphere and larger coordination environment of > 70 000 metal ions of 36 elements found in > 2000 structures of nucleic acids contained today in the PDB and NDB. MINAS is updated monthly with new structures and offers a multitude of search functions, e.g. the kind of metal ion, metal-ligand distance, innersphere and outersphere ligands defined by element or functional group, residue, experimental method, as well as PDB entry-related information. The results of each search can be saved individually for later use with so-called miniPDB files containing the respective metal ion together with the coordination environment within a 15 A radius. MINAS thus offers a unique way to explore the coordination geometries and ligands of metal ions together with the respective binding pockets in nucleic acids.

DOI: https://doi.org/10.1093/nar/gkr920

Originally published at:
Schnabl, J; Suter, P; Sigel, Roland K O (2012). MINAS-a database of Metal Ions in Nucleic AcidS. Nucleic Acids Research, 40(D1):D434-D438.
DOI: https://doi.org/10.1093/nar/gkr920

D434–D438 *Nucleic Acids Research, 2012, Vol. 40, Database issue* *Published online 16 November 2011*
doi:10.1093/nar/gkr920

MINAS—a database of *M*etal *I*ons in *N*ucleic *A*cid*S*

Joachim Schnabl, Pascal Suter and Roland K. O. Sigel*

Institute of Inorganic Chemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Received August 20, 2011; Revised October 5, 2011; Accepted October 8, 2011

ABSTRACT

Correctly folded into the respective native 3D structure, RNA and DNA are responsible for uncountable key functions in any viable organism. In order to exert their function, metal ion cofactors are closely involved in folding, structure formation and, e.g. in ribozymes, also the catalytic mechanism. The database MINAS, *M*etal *I*ons in *N*ucleic *A*cid*S* (http://www.minas.uzh.ch), compiles the detailed information on innersphere, outersphere and larger coordination environment of >70 000 metal ions of 36 elements found in >2000 structures of nucleic acids contained today in the PDB and NDB. MINAS is updated monthly with new structures and offers a multitude of search functions, e.g. the kind of metal ion, metal-ligand distance, innersphere and outersphere ligands defined by element or functional group, residue, experimental method, as well as PDB entry-related information. The results of each search can be saved individually for later use with so-called miniPDB files containing the respective metal ion together with the coordination environment within a 15 Å radius. MINAS thus offers a unique way to explore the coordination geometries and ligands of metal ions together with the respective binding pockets in nucleic acids.

INTRODUCTION

RNA is not only an information carrier for the primary protein structure, but is also actively involved in numerous processes within living cells (1). The RNA world hypothesis suggests that at early stages of evolution solely RNA acted as both information carrier and cellular catalyst (2). The existence of ribozymes discovered in the beginning of the 1980s (3,4) supports this hypothesis and ever since interest in native RNA structures has grown. Today, RNA molecules with manifold tasks are known. For example, so-called riboswitches are aptameric regions of mRNAs involved in self-regulation of genes in metabolic pathways (5). All RNA functions depend on a correct fold. Since nucleic acids contain a polyanionic phosphate-sugar backbone, this electrostatic repulsion has to be overcome to allow a close neighborhood of the negatively charged sites. Alkaline and earth alkaline metal ions are ideal for general charge screening due to their high abundance, their localized charge, and relatively weak general affinity towards any ligand. Aside from the negatively charged phosphate oxygens, each nucleotide comprises several potential coordinating atoms all of which may be important for 3D structure and catalysis. If these coordinating atoms are arranged ideally, metal ions can be recognized specifically and tightly bound through an intricate network of innersphere and outersphere interactions. The 2006 established MeRNA database categorizes eight such metal ion binding motifs together with their inner coordination environment within 6 Å identified in 389 PDB files (6). Specifically, bound metal ions are most crucial for structure and function of complex nucleic acid structures in general (7,8). For example, the necessity of divalent cations becomes obvious by the fact that the Mg^{2+} concentration can be reduced from ~50 mM to physiological 3 mM in the group II intron ribozyme Sc.ai5γ upon addition of the naturally associated protein Mss116, but cannot be completely replaced by other factors (9). A detailed knowledge on the coordination environment and the preferences of each metal ion for specific binding atoms is thus essential to understand structure and function of nucleic acids.

The most abundant divalent metal ion in living cells, Mg^{2+} (10), has the same number of electrons as H_2O and Na^+ and is mostly spectroscopically silent. Consequently, it is difficult to localize Mg^{2+} and to differentiate single Mg^{2+} ions from H_2O or Na^+ by X-ray, NMR, or other spectroscopic methods. High resolution X-Ray data are essential together with a correct interpretation of the electron density maps to assign the defined places of Mg^{2+} and other metal ions. A detailed knowledge on the most preferred coordinating atoms would thus be of great help to identify possible Mg^{2+} ions in electron density maps. Of the >5500 structures of nucleic acids deposited in the Protein Data Bank (PDB) (11,12), more than half contain metal ions that are bound to nucleic acids. Until today, there has been no possibility

*To whom correspondence should be addressed. Tel: +41 44 635 4652; Fax: +41 44 635 6802; Email: roland.sigel@aci.uzh.ch

to directly explore all these metal ions with regard to their first *and* second shell ligand environment, except for going through these PDB files individually.

The here described database MINAS, *M*etal *I*ons in *N*ucleic *A*cid*S* (http://www.minas.uzh.ch), now enables the detailed search for information on all metal ions and their wider coordination environment stored in the PDB (currently >2000 structures). MINAS is monthly updated and offers a large variety of intuitive search functions to explore all metal ions and their first and second shell coordination sphere identified in nucleic acid structures deposited in the PDB. The matching metal ions and ligands can then be viewed and downloaded as tabular data. Additionally, the binding pocket of the metal ion is provided as a section of the original PDB file displaying the macromolecular environment within a radius of 15 Å around the central metal ion. These so-called miniPDB files can be viewed directly in the web browser (Jmol, Java-plugin required) or downloaded for local display in a PDB viewer of choice. As of July 2011, the MINAS database contains over 70 000 metal ions of 36 distinct elements. The majority is accounted for by Mg^{2+} with over 50 000 entries being the most abundant metal ion. Mg^{2+} is followed by Na^{+} (more than 6000) and Sr^{2+} (more than 3000). With an almost exponential growth rate of the PDB, MINAS grows and more information on metal ions becomes available.

AIMS AND SCOPE OF THE MINAS DATABASE

Rising interest in metal binding sites in RNA requires that the existing wealth of information is readily accessible. To extend the knowledge about binding properties of metal ions and enable a direct and rapid access we created the database of MINAS (http://www.minas.uzh.ch). MINAS contains the full structural information of all nucleic acid–metal ion binding sites stored in the PDB. This centralized information allows now a detailed search of metal ion coordination sites in RNA, DNA and PNA: (i) The kind of metal ion can be specified along with multiple innersphere or outersphere ligands; (ii) These ligands can be defined by element, residue, distance to the metal ion and ligand relations; (iii) Apart from the coordination sphere, a variety of structurally related search criteria such as experimental methods can be chosen; (iv) PDB-related search terms like authors or keywords can be used to filter results. The retrieved information can then be used to understand the specificity of metal ions binding to nucleic acids. In addition, comparison with other nucleic acid molecules can accelerate the process of solving new 3D structures of nucleic acids containing metals and locate metal ions more exact in the electron density map.

MeRNA and MINAS have a different focus and yield different information. The MeRNA database (http://merna.lbl.gov) focuses on the eight known binding motifs derived from 389 RNA containing PDB files (6). MINAS instead lists all nucleic acid bound metal ions contained in the PDB and focuses on the ligands and their combination to build up a complete coordination sphere. MINAS will prove very useful to identify possible further binding motifs of metal ions in nucleic acids in general.

STRUCTURE OF THE DATABASE

All structure files of the PDB containing nucleic acids (RNA, DNA and PNA) form the basis of MINAS. These files were downloaded and screened for metal ions. Each metal ion was then screened for potential ligands (i.e. the elements nitrogen and oxygen) found in an appropriate distance. For innersphere ligands, the cutoff distance was set to 2.5 Å. Outersphere ligands were defined by a maximum distance of 3.2 Å to its respective innersphere ligand corresponding to the maximum length of a typical hydrogen bond. Hence, a sphere with a radius of 5.7 Å around each metal ion was defined as the coordination environment. The coordinates of each metal ion together with those of the potential ligands were then stored in the MINAS database along with the metadata from the PDB files such as experimental data and authors. In order to enable the convenient visualization of each metal ion binding site, the surrounding binding pocket within a radius of 15 Å around the metal ion is stored as a so-called miniPDB file. On a monthly basis, the PDB is searched for newly deposited PDB files containing nucleic acids to keep the MINAS database up-to-date.

ACCESS TO THE MINAS DATABASE

The MINAS database can be easily accessed at http://www.minas.uzh.ch. Access and full search options are free and unrestricted. The user obtains the full output list including shortcuts to the miniPDB files, which can be downloaded as *.csv files. In addition, the user can also register to obtain a login and password (registration is free). This allows the user to save each query and its output on the MINAS server for later retrieval, analysis and comparison.

SEARCH AND OUTPUT OF MINAS

The search portal of MINAS database can be found on http://www.minas.uzh.ch/search (Supplementary Figure S1). The query (unless defined differently by the user) automatically searches the full database containing more than 70 000 metal ions in >2000 PDB files of nucleic acids. The Query Builder is subdivided into three sections, which are briefly explained below. For a full documentation of the MINAS search functions, please refer to the Supplementary Information or the regularly updated help menu on the MINAS webpage:

(i) The top part of the Query Builder allows the user to define the search (Figure 1, Supplementary Figures S2–S8). On the top level, the metal ion of interest is set (Supplementary Figure S2), and in the middle level, PDB-related information can be specified, like PDB ID, experimental method, authors and journal, and the type of macromolecule (Supplementary Figures S3–S6). In the third level,

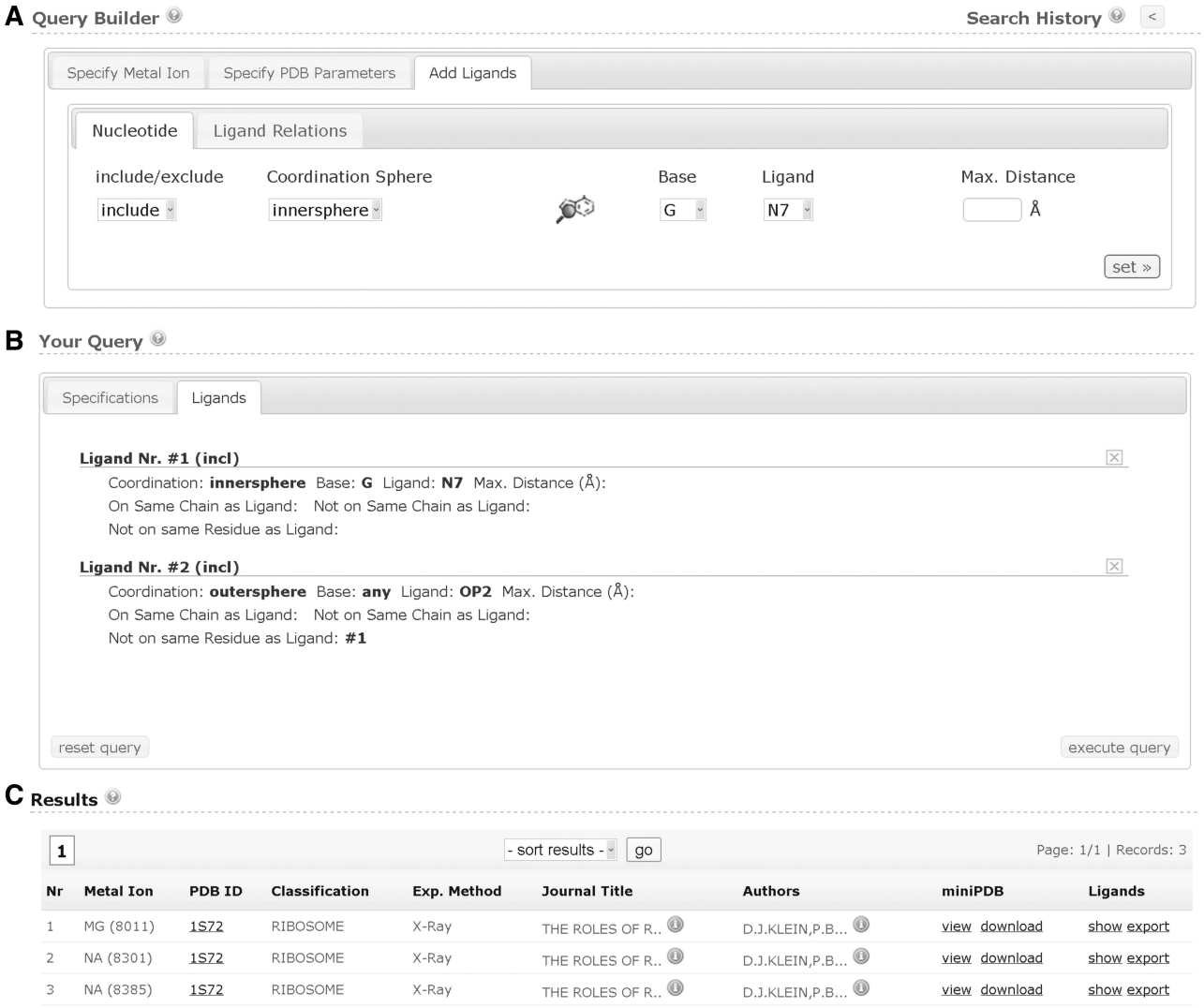

Figure 1. Screenshot of the query and results from http://www.minas.uzh.ch/search. (**A**) The *Query Builder* allows the specification of metal ions, the PDB-related data as well as the detailed choice of (multiple) ligands. (**B**) The section *Your Query* summarizes the specified parameters and choices for the query. (**C**) In the *Results* section, the metal ions found according to the set parameters are listed together the with the number of total hits in the top right corner. Links to the parent PDB ID, the miniPDB for viewing in a Java-enabled web browser as well as to a list of all identified ligands are provided (see also Figure 2).

the ligands in the first and second shell coordination sphere can specified (or excluded) (Figure 1A, Supplementary Figures S7 and S8). Ligands can be any coordinating atom of a nucleotide, water or simply the element nitrogen or oxygen. Also the relation between two coordinating atoms can be defined, e.g. on the same residue or two nucleotides apart. Any search definition needs to be confirmed with 'set'.

(ii) In the middle section, called Your Query, the search parameters as defined by the user in the Query Builder, are summarized (Figure 1B, Supplementary Figures S9 and S10). This section is intended to provide the user with an overview on the search, which is important especially for more complex multi-ligand searches. Once, the search is defined as intended, the search is executed by the 'execute query' button.

(iii) The last section on the bottom shows the results of the query (Figure 1C, Supplementary Figure S11). In the top right corner of this section, the total number of metal ions matching the user-specified query is given. The metal ions are listed below together with the PDB ID as well as the method and publication information. In column eight, the miniPBD can be viewed directly with Jmol (13) (when using a Java-enabled web browser; Figure 2A, Supplementary Figure S12) or downloaded for local viewing. In column nine, all potential ligands within a distance of 5.7 Å from the metal ion can be viewed (Figure 2B, Supplementary Figure S13) and/or downloaded as *.csv file. Previous searches and their results can be viewed with the Search History function (Supplementary Figure S14). If registered, the query results can also be saved online for later retrieval (Supplementary Figures S15 and S16).

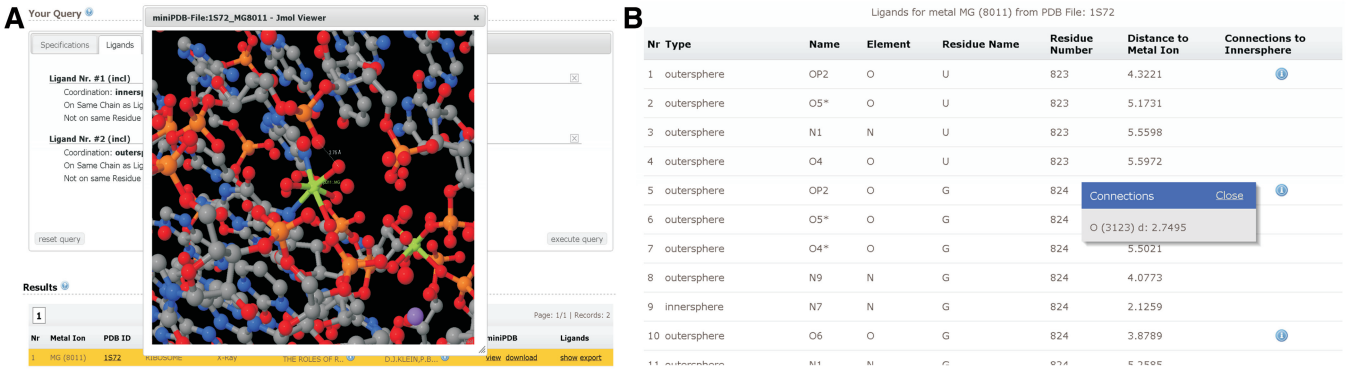

Figure 2. Screenshot of the detailed results from http://www.minas.uzh.ch/search. (**A**) View of a miniPDB file as provided after clicking on 'view'. A Jmol (13) frame displays the interactive 3D representation of the metal ion and its binding pocket. (**B**) A click on 'show' opens a table with all ligands listed. Here, details on the ligands from the PDB files as well as their direct distance to the metal ions are shown. Those outersphere ligands, which have a direct connection to an innersphere ligand display a blue info symbol in the column 'Connections to Innersphere'. This information box shows the element and residue number of the innersphere ligand as well as the distance between innersphere and outersphere ligands.

CONCLUSIONS

The MINAS database comprises as of July 2011 36 metallic elements counting more than 70 metal ions in total. Mg^{2+} is by far the most abundant metal in MINAS with over 50 000 entries coordinating to nucleic acids. Second to Mg^{2+} is Na^{+} with more than 6000 entries in the MINAS database and third is Sr^{2+} with still more than 3000 metals found. On the other end are Li^{+} and Lu^{3+} (1 entry each), as well as Sb^{3+} and Yb^{3+} (2 entries each). It is obvious that with increasing numbers of submitted macromolecular structures to the PDB each year (12), the number of metal ions increases with the same pace and the MINAS database will cover an even wider range of metal ions in nucleic acids.

Its sister database MeRNA (Metals in RNA) (6), established 2006 by Brenner and coworkers, focuses on binding motifs and thus has a different aim compared to MINAS. MeRNA does not include DNA and PNA but provides detailed information on the eight known types of metal ion binding pockets in RNA. The entries taken from 389 PDB files are classified together with their ligands making no distinction between inner- and outersphere. MINAS has a wider information and different focus being more ligand oriented. MINAS makes no preset restrictions, but instead includes *all* metal ions found in nucleic acids structures within the PDB together with their coordination environment. The stored structural data in a 15 Å radius around the metal ions allows for a detailed search and comparison of inner and outersphere ligands. The provided search functions allow the user to define any possible combination of coordinating ligands and parameters. MINAS is thus an ideal complementation to other databases like MeRNA (6) or SCOR (14). Further related databases are the Solvation Web Server (SwS) (15), which provides detailed information on the first solvation shell of nucleotides, and AMIGOS II (16), which allows the definition of and search for RNA structural motifs. In combination with those databases and their search functions, MINAS will provide new ways to investigate RNA structures, metal ion binding motifs and the role of metal ions in nucleic acid structure and function.

It is obvious that the data compiled in MINAS can only be as good as the initial structure, e.g. the resolution of an X-ray structure. In many cases, the resolution does not allow for a rigorous search for metal ions and sometimes, the metal ions have also not been refined. In the future as the PDB grows and the structures generally improve in resolution with the always better technical possibilities, also the quality of the data compiled in MINAS will rise. This will lead to more results of a given query and thus statistically more meaningful data. Consequently, our understanding of metal ion binding will increase, e.g. by finding new classes of metal ion binding pockets in RNA.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online: Supplementary Information, Supplementary Figures 1–16.

ACKNOWLEDGEMENTS

We are very grateful to all users of the MINAS database and especially to those that report problems and help us improve our database. Financial support from the various funding sources (UZH, SNSF, SBF, COST, ERC) is gratefully acknowledged.

FUNDING

University of Zurich; the Swiss National Science Foundation (200021_124834 to R.K.O.S.); the Swiss State Secretariat for Education and Research (COST Action D39, C07.0129 to R.K.O.S.); European Research Council (ERC Starting Grant MIRNA 259092 to R.K.O.S.). Funding for open access charge: University of Zurich.

Conflict of interest statement. None declared.

REFERENCES

1. Gesteland,R.F., Cech,T.R. and Atkins,J.F. (2006) *The RNA World*, 3rd edn. Cold Spring Harbor Laboratory, New York.
2. Gilbert,W. (1986) Origin of life: the RNA world. *Nature*, **319**, 618.
3. Guerrier-Takada,C., Gardiner,K., Marsh,T., Pace,N. and Altman,S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35**, 849–857.
4. Kruger,K., Grabowski,P.J., Zaug,A.J., Sands,J., Gottschling,D.E. and Cech,T.R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell*, **31**, 147–157.
5. Roth,A. and Breaker,R.R. (2009) The structural and functional diversity of metabolite-binding riboswitches. *Annu. Rev. Biochem.*, **78**, 305–334.
6. Stefan,L.R., Zhang,R., Levitan,A.G., Hendrix,D.K., Brenner,S.E. and Holbrook,S.R. (2006) MeRNA: a database of metal ion binding sites in RNA structures. *Nucleic Acids Res.*, **34**, D131–D134.
7. Schnabl,J. and Sigel,R.K.O. (2010) Controlling ribozyme activity by metal ions. *Curr. Opin. Chem. Biol.*, **14**, 269–275.
8. Sigel,R.K.O. and Pyle,A.M. (2007) Alternative roles for metal ions in enzyme catalysis and the implication for ribozyme chemistry. *Chem. Rev.*, **107**, 97–113.
9. Solem,A., Zingler,N. and Pyle,A.M. (2006) A DEAD protein that activates intron self-splicing without unwinding RNA. *Mol. Cell*, **24**, 611–617.
10. Pechlaner,M. and Sigel,R.K.O. (2012) Characterization of metal ion-nucleic acid interactions in solution. *Met. Ions Life Sci.*, **10**, 1–42.
11. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
12. Rose,P.W., Beran,B., Bi,C., Bluhm,W.F., Dimitropoulos,D., Goodsell,D.S., Prlić,A., Quesada,M., Quinn,G.B., Westbrook,J.D. *et al.* (2011) The RCSB Protein Data Bank: redesigned web site and web services. *Nucleic Acids Res*, **39**, D392–D401.
13. Herráez,A. (2006) Biomolecules in the computer: Jmol to the rescue. *Biochem. Mol. Biol. Edu.*, **34**, 255–261.
14. Tamura,M., Hendrix,D.K., Klosterman,P.S., Schimmelman,N.R.B., Brenner,S.E. and Holbrook,S.R. (2004) SCOR: Structural classification of RNA, version 2.0. *Nucleic Acids Res*, **32**, D182–D184.
15. Auffinger,P. and Hashem,Y. (2007) SwS: a solvation web service for nucleic acids. *Bioinformatics*, **23**, 1035–1037.
16. Wadley,L.M., Keating,K.S., Duarte,C.M. and Pyle,A.M. (2007) Evaluating and learning from RNA pseudotorsional space: quantitative validation of a reduced representation for RNA structure. *J. Mol. Biol.*, **372**, 942–957.